# WEAK CONVERGENCE OF KOLMOGOROV-SMIRNOV TYPE STATISTICS

by

Constance L. Wood
Cornell University

1. Introduction. It is often of interest to test whether a random sample has a distribution from a specified scale and translation family, with the particular scale and translation parameters unknown. If the values of these parameters were known, the data could be reduced to values having a distribution which is independent of the said parameters (standardized cumulative distribution function). A test could then be based on the well-known result that, under the null hypothesis, the (normalized) stochastic process, based on the difference of the sample cumulative distribution function (c.d.f.) of the transformed random variables and the standardized c.d.f., converges weakly to the tied-down Wiener process (Billingsley (1968), Thm. 13.1). Also, under a fixed alternative hypothesis, the limiting behavior of Kolmogorov-Smirnov type statistics, which are functionals of the empirical process, have been characterized by Raghavachari (1973) in terms of functionals of the limiting Wiener Process. However, if estimates of the parameters of the distribution are substituted for the actual parameters in performing the transformation, these results are no longer valid.

While considering the limiting null distribution of statistics which are functionals of this modified empirical stochastic process, many authors have extensively studied the finite-dimensional distributions of the limiting process. Darling (1955), having first derived the finite-dimensional distributions, found the asymptotic characteristic function of the Cramér - von Mises statistic when the null family of distributions admits either a weakly unbiased maximum likelihood or an efficient (in the sense of Cramér) estimator ($\hat{\theta}_n$) of the parameter ($\theta$), which is

essentially a sum of independently and identically distributed (i.i.d.) random variables (r.v.); i.e., $n^{\frac{1}{2}}(\hat{\theta}_n - \theta) = \sum_{i=1}^{n} \ell(X_i,\theta) + \epsilon_n$, $\epsilon_n = o_p(1)$. Restricting attention to tests for normality, Kac, Kiefer, and Wolfowitz (1955) investigated the same problem for the case where both the mean and variance were unknown. Durbin (1973), extending Darling's approach to multiparameter families, obtained the limiting process under both the null hypothesis and a sequence of parametric alternatives from the null family of distributions which converge to the null at the rate $n^{-\frac{1}{2}}$.

The weak convergence of the modified empirical process under only weak regularity conditions on the standardized c.d.f. and asymptotic normality of the estimates is explored in Section 2 of this paper. These results are used in the third section as basis for a Monte Carlo study of the distribution of Kolmogorov-Smirnov type statistics under the null hypothesis. The fourth section contains the asymptotic properties of the statistics mentioned above, under a fixed alternative hypothesis.

2. Asymptotic Distribution of an Empirical Stochastic Process under $H_o$. In connection with independent observations $X_1, X_2, \cdots, X_n$ on a probability space $(\Omega,A,P)$ and having distribution F, suppose that we are interested in the null hypothesis

$$(2.1) \qquad H_o: F(x) = H\left(\frac{x - \alpha}{\beta}\right) \text{ for some choice of } (\alpha,\beta), \text{ say } (\alpha_H,\beta_H).$$

Here we assume that H is specified but not necessarily $\alpha_H$ or $\beta_H$. Thus $H_o$ is a composite hypothesis.

Under $H_o$, the random functions

$$(2.2) \qquad H_n(t) = n^{-1} \sum_{i=1}^{n} I\left[H\left(\frac{X_i - \alpha_H}{\beta_H}\right) \le t\right], \quad 0 \le t \le 1,$$

and

$$W_n(t) = \sqrt{n}\,[H_n(t) - t], \quad 0 \le t \le 1, \tag{2.3}$$

are well defined and, by a well-known result (Billingsley (1968), Thm. 13.1),

$$W_n \xrightarrow{\mathcal{L}} W^o, \quad \text{in } D[0,1], \tag{2.4}$$

where $W^o$ is the "tied-down Wiener Process" on $[0,1]$.

If $(\alpha_H,\beta_H)$ were specified by $H_o$, then $W_n(\cdot)$ could be written down as a function of the observations and a test of $H_o$ carried out in terms of $W_n(\cdot)$ and the asymptotic result (2.4). However, $(\alpha_H,\beta_H)$ not being specified, we consider an analogous approach replacing $(\alpha_H,\beta_H)$ by estimates $(\hat{\alpha}_n,\hat{\beta}_n)$ consistent under $H_o$. (More specific assumptions on $(\hat{\alpha}_n,\hat{\beta}_n)$ will be given in Lemma 2.2.) Now define

$$Y_{ni} = (X_i - \hat{\alpha}_n)/\hat{\beta}_n, \quad 1 \le i \le n. \tag{2.5}$$

In analogy with $H_n(\cdot)$ and $W_n(\cdot)$, let

$$G_n(t) = n^{-1} \sum_{i=1}^{n} I\Big[H(Y_{ni}) \le t\Big], \quad 0 \le t \le 1, \tag{2.6}$$

and

$$V_n(t) = \sqrt{n}\,\Big(G_n(t) - t\Big), \qquad 0 \le t \le 1. \tag{2.7}$$

First note that for $0 \le t \le 1$,

$$G_n(t) = n^{-1} \sum_{i=1}^{n} I\Big[H(Y_{ni}) \le t\Big]$$

$$= H_n\Big\{H\Big[(\hat{\beta}_n/\beta_H)H^{-1}(t) + (\hat{\alpha}_n - \alpha_H)/\beta_H\Big]\Big\}.$$

Thus

$$V_n(t) = \sqrt{n}\left[H_n(\phi_n(t) - t)\right], \quad 0 \le t \le 1, \tag{2.8}$$

where

$$\phi_n(t) = H\left[(\hat{\beta}_n/\beta_H)H^{-1}(t) + (\hat{\alpha}_n - \alpha_H)/\beta_H\right], \quad 0 \le t \le 1. \tag{2.9}$$

Note that $\phi_n(t)$ is increasing in t, and hence is a "random change of time" in the sense of Billingsley (1968), p. 144.

With this notation,

$$V_n(t) = \sqrt{n}\left[H_n[\phi_n(t)] - \phi_n^{-1}[\phi_n(t)]\right], \quad 0 \le t \le 1, \tag{2.10}$$

$$= \Delta_n \circ \phi_n(t),$$

where

$$\Delta_n(t) = \sqrt{n}\left[H_n(t) - \phi_n^{-1}(t)\right], \quad 0 \le t \le 1. \tag{2.11}$$

LEMMA 2.1. Assume $H''(x)$ is bounded and $xH'(x) \to 0$ as $|x| \to \infty$. If $\sqrt{n}\,\hat{\theta}_{n1}$ and $\sqrt{n}\,(\hat{\theta}_{n2} - 1)$ are each $o_p(1)$, $n \to \infty$, then

$$\sqrt{n} \sup_x \left| H(\hat{\theta}_{n2}x + \hat{\theta}_{n1}) - H(X) - H'(x)\left[(\hat{\theta}_{n2} - 1)x + \hat{\theta}_{n1}\right] \right| = o_p(1), \; n \to \infty. \tag{2.12}$$

COROLLARY 2.1. Under the conditions of LEMMA 2.1,

$$\sqrt{n} \sup_x \left| H(\hat{\theta}_{n2}x + \hat{\theta}_{n1}) - H(x)\right| = o_p(1), \; n \to \infty. \tag{2.13}$$

PROOF. Clearly we have that

$$\sqrt{n} \sup_x H'(x)\left[(\hat{\theta}_{n2} - 1)x + \hat{\theta}_{n1}\right] = o_p(1), \; n \to \infty. \tag{2.14}$$

The result follows immediately from (2.12).

Q.E.D.

We will make the following definitions:

$$I(t) = t, \quad 0 \le t \le 1, \tag{2.15}$$

and

$$\Delta_n^*(t) = \sqrt{n}\left\{H_n(t) - t + H'\left(H^{-1}(t)\right)\left[H^{-1}(t)\left(\frac{\hat{\beta}_n - \beta_H}{\beta_H}\right) + \left(\frac{\hat{\alpha}_n - \alpha_H}{\beta_H}\right)\right]\right\}, \quad 0 \le t \le 1. \tag{2.16}$$

LEMMA 2.2. *If $H''(x)$ is bounded, $xH'(x) \to 0$ as $|x| \to \infty$, and both $\sqrt{n}\,(\hat{\alpha}_n - \alpha_H)$ and $\sqrt{n}\,(\hat{\beta}_n - \beta_H)$, $\beta_H > 0$, have non-degenerate limit laws, then*

$$\phi_n \xrightarrow{\mathcal{L}} I \quad , \quad \text{in } D[0,1], \tag{2.17}$$

and

$$\Delta_n - \Delta_n^* \xrightarrow{p} 0, \quad \text{in } D[0,1]. \tag{2.18}$$

Choose $0 \le t_1, t_2, \cdots, t_k \le 1$ and consider the following vector:

$$\underline{\eta}_n' = \left[W_n(t_1), \cdots, W_n(t_k), \sqrt{n}\,(\hat{\alpha}_n - \alpha_H)/\beta_H, \sqrt{n}\,(\hat{\beta}_n - \beta_H)/\beta_H\right]. \tag{2.19}$$

LEMMA 2.3. *If $H'(x)$ is a continuous function of $x$ and positive on the support of $H$, and there exists a positive definite matrix $\Lambda = (\lambda_{ij})$ such that $\underline{\eta}_n$ is AMN$(\underline{0}, \Lambda)$, then*

$$\Delta_n^* \xrightarrow{\mathcal{L}} V^0, \quad \text{in } D[0,1], \tag{2.20}$$

*where $V^0$ is the Gaussian process determined by*

$$E[V^0(t)] = 0, \quad 0 \le t \le 1, \tag{2.21}$$

*and, for $0 \le t_1, t_2 \le 1$,*

$$(2.22)\quad E[V^{o}(t_1)V^{o}(t_2)] = \min(t_1,t_2) - t_1t_2 + H'[H^{-1}(t_1)]H'[H^{-1}(t_2)]\lambda_{k+1,k+1}$$
$$+ H^{-1}(t_1)H'[H^{-1}(t_1)]H^{-1}(t_2)H'[H^{-1}(t_2)]\lambda_{k+2,k+2} + H'[H^{-1}(t_2)]\lambda_{1,k+1}$$
$$+ H'[H^{-1}(t_1)]\lambda_{2,k+1} + H^{-1}(t_1)H'[H^{-1}(t_1)]\lambda_{2,k+2}$$
$$+ H^{-1}(t_2)H'[H^{-1}(t_2)]\lambda_{1,k+2} + \left[H^{-1}(t_1)H'[H^{-1}(t_1)]H'[H^{-1}(t_2)]\right.$$
$$\left. + H^{-1}(t_2)H'[H^{-1}(t_2)]H'[H^{-1}(t_1)]\right]\lambda_{k+1,k+2}$$

THEOREM 2.1. Suppose that $\underline{\eta}_n$ is AMN($\underline{0},\Lambda$), where $\underline{\eta}_n$ is given in (2.19). Further suppose that $\sup_x H''(x) < \infty$, $\lim_{|x|\to\infty} xH'(x) = 0$, and $H'(x)$ is positive on the support of H. Then

$$(2.23)\quad \lim_{n\to\infty} \mathcal{L}[V_n|H_o] = \mathcal{L}[V^{o}], \quad \text{in } D[0,1],$$

where $V^{o}$ is the Gaussian Process given in LEMMA 2.3.

PROOF. First from Lemma 2.2 and Lemma 2.3, we have that

$$(2.24)\quad \phi_n \xrightarrow{p} I \quad , \quad \text{in } D[0,1],$$

$$(2.25)\quad \Delta_n^{*} - \Delta_n \xrightarrow{p} 0, \quad \text{in } D[0,1],$$

and

$$(2.26)\quad \Delta_n^{*} \xrightarrow{\mathcal{L}} V^{o} \quad , \quad \text{in } D[0,1].$$

Hence, by Theorem 4.1 of Billingsley (1968), we have that

$$(2.27)\quad \Delta_n \xrightarrow{\mathcal{L}} V^{o} \quad , \quad \text{in } D[0,1].$$

Thus, by Theorem 4.4 of Billingsley (1968),

$$(2.28)\quad (\Delta_n,\phi_n) \xrightarrow{\mathcal{L}} (V^{o},I), \quad \text{in } D[0,1].$$

Consequently, since $P\left[V^{\circ} \in C[0,1]\right] = 1$, we have, by Billingsley (1968), Section 17.1, that

$$\Delta_n \circ \phi_n \xrightarrow{\mathcal{L}} V^{\circ} \circ I = V^{\circ}, \quad \text{in } D[0,1]. \tag{2.29}$$

Q.E.D.

3. Monte Carlo Results for Kolmogorov-Smirnov Type Statistics Under the Null Hypothesis. We want to consider statistics which are functionals of the modified empirical process. In particular, the one-sample statistics to be examined are as follows:

(i) One-sided Kolmogorov-Smirnov statistics

$$D_n^{+} = \sup_{0 \le t \le 1} \sqrt{n}\,[G_n(t) - t] \tag{3.1}$$

and

$$D_n^{-} = \inf_{0 \le t \le 1} \sqrt{n}\,[G_n(t) - t]. \tag{3.2}$$

(ii) Kolmogorov-Smirnov statistic

$$D_n = \sup_{0 \le t \le 1} \sqrt{n}\,|G_n(t) - t| \tag{3.3}$$

and

(iii) Kupier statistic

$$D_n^{\pm} = (D_n^{+} - D_n^{-}). \tag{3.4}$$

Our Monte Carlo results are based on the following well-known theorem.

THEOREM 3.1. Under the conditions of THEOREM 2.1,

$$\text{(3.5)} \qquad \text{(i)} \quad \lim_{n\to\infty} \mathcal{L}[D_n^+ | H_o] = \mathcal{L}\Big[\sup_{0\le t\le 1} [V^o(t)]\Big]$$

$$\text{(3.6)} \qquad \text{(ii)} \quad \lim_{n\to\infty} \mathcal{L}[D_n^- | H_o] = \mathcal{L}\Big[\inf_{0\le t\le 1} [V^o(t)]\Big]$$

$$\text{(3.7)} \qquad \text{(iii)} \quad \lim_{n\to\infty} \mathcal{L}[D_n | H_o] = \mathcal{L}\Big[\sup_{0\le t\le 1} |V^o(t)|\Big]$$

and

$$\text{(3.8)} \qquad \text{(iv)} \quad \lim_{n\to\infty} \mathcal{L}[D_n^{\pm} | H_o] = \mathcal{L}\Big[\sup_{0\le t\le 1} V^o(t) - \inf_{0\le t\le 1} V^o(t)\Big],$$

where $V^o(\cdot)$ is the Gaussian Process given in LEMMA 2.3.

The procedure was to approximate the Gaussian Process $V^o(\cdot)$ by its finite-dimensional distributions corresponding to a division of the unit interval into equal sub-intervals. One thousand multivariate normally distributed vectors were then generated having covariance function corresponding to the null distribution and estimates of the parameters. Each of the functionals were then applied to the vector to obtain an approximation to the functionals of the Gaussian Process.

Here we will only consider the Kolmogorov-Smirnov statistics for testing normality. The unknown parameters will be estimated by the sample mean and the sample variance. The limiting Gaussian Process ($V^o$) was approximated by its finite-dimensional distributions taken at spacings of 1/30, 1/60, and 1/120. Quantiles and moments of the sampling distributions, generated for the Kolmogorov-Smirnov type statistics, are given in the tables below:

TABLE 3.1

SAMPLING DISTRIBUTIONS OF APPROXIMATE KOLMOGOROV-SMIRNOV TYPE STATISTICS FOR TESTING NORMALITY WITH MEAN AND VARIANCE ESTIMATED $(\bar{x}, s^2)$*

| Quantile | Statistic | | | |
|---|---|---|---|---|
| | $D_n^+$ | $D_n^-$ | $D_n$ | $D_n^{\pm}$ |
| .010 | 0.249 | -0.926 | 0.327 | 0.614 |
| .025 | 0.268 | -0.841 | 0.351 | 0.670 |
| .050 | 0.293 | -0.768 | 0.381 | 0.704 |
| .100 | 0.319 | -0.698 | 0.409 | 0.770 |
| .250 | 0.383 | -0.583 | 0.469 | 0.880 |
| .500 | 0.469 | -0.478 | 0.555 | 1.033 |
| .750 | 0.573 | -0.394 | 0.659 | 1.221 |
| .900 | 0.685 | -0.333 | 0.752 | 1.414 |
| .950 | 0.746 | -0.300 | 0.835 | 1.576 |
| .975 | 0.826 | -0.272 | 0.910 | 1.698 |
| .990 | 0.926 | -0.237 | 0.991 | 1.853 |

* Based on a grid of 120 sub-intervals.

TABLE 3.2

EXACT PROBABILITIES OF DEVIATIONS OF EMPIRICAL C.D.F. FROM TRUE C.D.F.

$$\Pr\left[\sup_x |F_n(x) - F(x)| > r/\sqrt{n}\right]$$

| N | $r/\sqrt{n}$ | Probability |
|---|---|---|
| 500 | 0.1000 | 0.00000008 |
| | 0.0100 | 0.00835507 |
| | 0.0050 | 0.08335292 |
| | 0.0025 | 0.28746880 |
| 1000 | 0.0100 | 0.00085645 |
| | 0.0050 | 0.00757943 |
| | 0.0025 | 0.08259189 |

TABLE 3.3

SAMPLING DISTRIBUTION OF APPROXIMATE KOLMOGOROV-SMIRNOV STATISTIC FOR TESTING NORMALITY WITH MEAN AND VARIANCE ESTIMATED $(\bar{x}, s^2)$

| Quantile | Number of Sub-Intervals | | | | | |
|---|---|---|---|---|---|---|
| | 20 | 30 | 40 | 60 | 80 | 120 |
| 0.010 | 0.237 | 0.277 | 0.306 | 0.319 | 0.314 | 0.327 |
| 0.050 | 0.298 | 0.330 | 0.357 | 0.358 | 0.376 | 0.381 |
| 0.100 | 0.339 | 0.362 | 0.384 | 0.389 | 0.403 | 0.409 |
| 0.250 | 0.403 | 0.423 | 0.443 | 0.451 | 0.467 | 0.469 |
| 0.500 | 0.494 | 0.506 | 0.527 | 0.535 | 0.556 | 0.555 |
| 0.750 | 0.598 | 0.601 | 0.638 | 0.633 | 0.667 | 0.659 |
| 0.900 | 0.708 | 0.694 | 0.749 | 0.735 | 0.780 | 0.752 |
| 0.950 | 0.781 | 0.790 | 0.821 | 0.823 | 0.849 | 0.835 |
| 0.990 | 0.896 | 0.945 | 0.978 | 0.968 | 0.982 | 0.991 |

TABLE 3.4

SAMPLE MEAN AND VARIANCE OF APPROXIMATE KOLMOGOROV-SMIRNOV STATISTIC FOR TESTING NORMALITY WITH MEAN AND VARIANCE ESTIMATED $(\bar{x}, s^2)$

| | Number of Sub-Intervals | | | | | |
|---|---|---|---|---|---|---|
| | 20 | 30 | 40 | 60 | 80 | 120 |
| SAMPLE MEAN | 0.510 | 0.521 | 0.550 | 0.554 | 0.576 | 0.574 |
| SAMPLE VARIANCE | 0.022 | 0.020 | 0.022 | 0.020 | 0.022 | 0.020 |